Progress Energy

Timika Shafeek-Horton
Deputy General Counsel
550 South Tryon Street
Charlotte, NC 28202

Mailing Address:
DEC 45A/PO Box 1321
Charlotte, NC 28201
704 382 6373 Direct
980 373 8534 Fax

Email Timika.Shafeek-Horton@duke-energy.com

October 1, 2012

The Honorable Jocelyn G. Boyd
Chief Clerk/Administrator
Public Service Commission of South Carolina
101 Executive Center Drive (29210)
Post Office Drawer 11649
Columbia, South Carolina 29211

Re: **Docket No. 2008-251-E**

Dear Ms. Boyd:

Pursuant to the Commission's May 6, 2009 Directive in Docket No. 2008-251-E approving its cost recovery mechanism, Progress Energy Carolinas, Inc. submits the attached report summarizing the results of the 2011 Program Year evaluation, measurement and verification (EM&V) for its Demand Response Automation Program. Progress Energy Carolinas, Inc. is currently evaluating the recommendations provided in the EM&V report.

Sincerely,

Timika Shafeek-Horton
Deputy General Counsel

Attachment
cc: Courtney Edwards
John Flitter

NAVIGANT

2011 EM&V Report for the Commercial, Industrial and Governmental Demand Response Automation (DRA) Program

Presented to:

Progress Energy Carolinas

Prepared by:

Navigant Consulting

September 28, 2012

NAVIGANT

Prepared for:

Progress Energy Carolinas
Raleigh, North Carolina

Presented by

Stuart Schare
Director

Navigant Consulting, Inc.
1375 Walnut Street
Suite 200
Boulder, CO 80302
phone 303.728.2500
fax 303.728.2501

www.navigantconsulting.com

Primary contributing author:

Peter Steele-Mosey

NAVIGANT

Table of Contents

Executive Summary

The Commercial, Industrial, and Governmental Demand Response Automation (DRA) program is part of the portfolio of demand side management and energy efficiency (DSM/EE) programs initiated by Progress Energy Carolinas (PEC) beginning in 2009. DRA offers participating companies and agencies a financial incentive to reduce their electricity consumption when called upon by PEC. This report covers evaluation, measurement, and verification (EM&V) activities for the second year of DRA, Program Year 2011 (PY2011).

This EM&V report is intended to verify program impacts as per the requirements established by the North Carolina Utilities Commission and the Public Service Commission of South Carolina. Major objectives of the evaluation were as follows:

» Verify the demand reduction calculated by PEC's method of baseline estimation as described in the *Demand Response Automation Rider DRA-1A (North Carolina), and DRA-2A (South Carolina)* filed by PEC in June of 2010.[1]

» Determine the most accurate baseline calculation method for the largest industrial participant's 16 different meters participating in the program.

» Produce a set of verified program impacts by customer and for the program as a whole, using the most accurate baseline method identified for the largest industrial participant (determined in PY2011) and for the balance of participants' meters (based on PY2010's analysis).

Program Summary

The DRA program offers participating companies and agencies a financial incentive to reduce their electricity consumption for up to eight hours at a time on a few peak summer days. Under the program, PEC's technology vendor (Comverge) installs two-way communications equipment to remotely monitor and record interval loads. Customer load curtailments are commonly provided through use of onsite generation or from shut down of manufacturing processes, although curtailments are also possible through modifications in the use of HVAC systems, lighting, and other building loads.

Fourteen customers actively participated in PEC's DRA program in the summer of 2011, representing 31 unique sites and 39 meters. Fourteen of the 39 meters are at commercial sites and five are at governmental sites (water treatment and detention facilities). Twenty meters are

---

[1] North Carolina Rider, DRA-1A: https://www.progress-energy.com/assets/www/docs/company/NCRiderDRA.pdf
South Carolina Rider, DRA-2A: https://www.progress-energy.com/assets/www/docs/business/scriderdra.pdf

at industrial sites, sixteen of which belong to a single manufacturing company. For brevity, the very large industrial participant (with 16 meters) is referred to in this report as the "VLIP".

An overview of the participating customers is presented in Table 1.

Table 1: Summary of Participating Companies and Agencies

Sector	Customer Type	Number of Participants	Number of Sites	Number of Meters
Commercial	Warehouse/Distribution	1	1	1
Industrial	Manufacturing	4	14	20
Governmental	Gov't Institution	1	2	2
Governmental	Water Treatment	3	3	3
Commercial	Grocery	2	8	8
Commercial	Data Center	1	1	3
Commercial	Office	1	1	1
Commercial	Hospital/Medical	1	1	1
	Total Program:	**14**	**31**	**39**

Source: PEC DRA program database

Impact Evaluation Methods

The method by which the evaluation team accomplished the principal goals set for the impact evaluation was comprised of three major components, as described below.

1. Replication of PEC Reported Impacts.

   The evaluation team used interval data for all participant meters and event data to calculate a baseline for each event and each participant meter. These baselines were all calculated using the PEC algorithm, which the company uses to report program impacts and calculate participant incentives for customer settlement purposes. This algorithm calculates the baseline as the average demand observed during the same hours as the event drawn from the first three prior non-excluded (i.e., non-holiday and non-event weekdays) qualifying days.

2. Estimation of Program Impacts for Planning Purposes

   In the evaluation of the DRA program completed in December 2011 (for PY2010), Navigant tested over a hundred different baselines, using both day-matching and regression techniques, to determine that which provided the most accurate estimate of participants' baselines. This testing showed that the most accurate method of baseline estimation was the use of an individual regression with a symmetric day-of adjustment. Note that although this was the most accurate method tested, it is unsuitable for settlement purposes as it requires an entire season's worth of data and is less reproducible and transparent than the current day-matching method.

3. Testing Accuracy of Baselines for DRA's Largest Industrial Participant

As noted above, 16 of the 39 meters participating in this program in PY2011 belonged to a single very large industrial (manufacturing) participant (the VLIP, accounting for 23% of reported load reductions). Given the high proportion of program demand impacts that were due to this single customer, PEC requested that Navigant repeat the testing exercise used for PY2010 to determine the most accurate method for estimating demand impacts for planning purposes, but this time with a separate, additional effort focused on the 16 meters belonging to the VLIP. As in PY2010, the evaluation team tested over a hundred different algorithms and models, including both day-matching algorithms (very much like the PEC algorithm) and regression models. Baselines for sixty non-event summer days on which the maximum temperature exceeded 85 degrees Fahrenheit were calculated and compared with actual demand during those times. The degree of difference between the demand predicted by the baseline and the actual demand was used to rank the different models and algorithms. On this basis, the most accurate was then used to estimate impacts.

Program Impact Findings

Three DRA events were called during summer 2011, involving 39 unique customer meters. The EM&V analysis found average load reductions of 13.4 MW per event (roughly 343 kW per meter or 946 per participant), or approximately 92% of the 14.6 MW figure reported[2] by PEC in its DRA program database (Table 2).

Table 2: Verified Load Reductions and EM&V Verification Rate

Load Reduction Category	July 12, 2011 (kW)	July 22, 2011 (kW)	Aug 08, 2011 (kW)	Avg. Total Reduction Over Three Events (kW)
Reported (PEC Database)	13,533	15,322	14,849	14,568
Verified	12,820	13,889	13,453	13,387
Realization Rate (Verified Reductions/Reported Reductions)	95%	91%	91%	92%

Source: PEC DRA program database and Navigant analysis

Note that the realization rate reported in the table above is the "verified realization rate", that is, the evaluation team's estimated demand impacts divided by the PEC reported settlement

[2] "Reported" impacts are those impacts calculated by PEC using the DRA baseline algorithm.

demand impacts. The other realization rate mentioned in this report is the "contracted realization rate" – the evaluation team's estimated demand impact divided by the participants' or participant's contracted demand reduction.

The evaluation team found in its analysis that the VLIP's demand was highly variable across many of its meters in the summer of 2011 – on some non-holiday weekdays demand was close to zero and on others in the range of hundreds of kilowatts. These volatile patterns of use mean that the estimated baselines and impacts for each of these meters individually tends to be less reliable than for the other meters and customers with a more consistent pattern of demand. However, the aggregate estimate of the VLIP's demand reduction impacts appear to be accurate, with estimated reductions for some individual meters likely overestimating true impacts and for others underestimating true impacts.

Recommendations

The evaluation team recommends a variety of discrete actions for improving the effectiveness of the DRA program in future years. These recommendations are based on insights gleaned from program and participant interval data analyzed for PY2011, and a comparison with the results in PY2010. Section 4.2 includes discussion of each of these recommendations.

Recommendation Topic	Recommended Actions
Improvements to Baseline Estimation	1. Modify the settlement baseline algorithm to exclude both low and high outlier days (not just low outlier days).
	2. Modify the settlement baseline algorithm to include a symmetric day-of adjustment.
	3. Continue to use a regression-derived baseline incorporating a 4-hour symmetric day-of adjustment for the estimation of system load impacts and for planning system resource needs.
Change to Event Notification Time	4. Consider leaving a smaller gap between event notification and event starting hours to improve the performance of the day-of load adjustment.
Participant Recruitment	5. Consider the consistency of potential participants' daily demand when targeting recruitment efforts and communicating with participants.
	6. Continue and enhance recruitment of new participants from the hospital sector.

Section 1. Introduction

The Commercial, Industrial, and Governmental Demand Response Automation program, also known as the CIG Demand Response Automation (DRA) program, is part of the portfolio of demand side management and energy efficiency (DSM/EE) programs initiated by Progress Energy Carolinas (PEC) beginning in 2009. DRA offers participating companies and agencies a financial incentive to reduce their electricity consumption for up to eight hours at a time on a few peak summer days. This report covers evaluation, measurement, and verification (EM&V) activities for the second year of DRA, Program Year 2011 (PY2011).

EM&V is a term adopted by PEC and refers generally to the assessment and quantification of the energy and peak demand impacts of an energy efficiency or demand response program. For DR, estimating reductions in peak demand is the primary objective, as energy impacts are generally negligible.

1.1 *Objectives of the Evaluation*

This EM&V report is intended to verify program impacts as per the requirements established by the North Carolina Utilities Commission and the Public Service Commission of South Carolina. Major objectives of the evaluation were as follows:

- Verify the demand reduction calculated by PEC's method of baseline estimation as described in the *Demand Response Automation Rider DRA-1A (North Carolina), and DRA-2A (South Carolina)* filed by PEC in June of 2010. [3]

- Determine the most accurate baseline calculation method for the largest industrial participant's 16 different meters participating in the problem.

- Produce a set of verified program impacts by customer and for the program as a whole, using the most accurate baseline method identified for the largest industrial participant (determined in PY2011) and for the balance of participants' meters (based on PY2010's analysis).

1.2 *Program Overview*

The DRA program was developed in response to PEC's determination that a curtailable load program would be a valuable resource for the company and an additional service offering for customers that would complement PEC's existing load curtailment riders. The program seeks to increase PEC's DR resources by improving customer receptiveness to curtailment programs

---

[3] North Carolina Rider, DRA-1A: https://www.progress-energy.com/assets/www/docs/company/NCRiderDRA.pdf
South Carolina Rider, DRA-2A: https://www.progress-energy.com/assets/www/docs/business/scriderdra.pdf

through increased awareness of load reduction potential and restructuring of the incentives and non-compliance charges used for current DR programs.

The program offers participating companies and agencies a financial incentive to reduce their electricity consumption for up to eight hours[4] at a time on a few peak summer days. Under the program, PEC's technology vendor (Comverge) installs two-way communications equipment to remotely monitor and record interval loads.

Eligibility. To qualify for the program, PEC commercial and industrial customers must have at least 200 kW of peak demand and be able to control 75 kW. Importantly, all industrial customers, and any commercial customers that use more than one million kWh per year, must also elect to forego the opportunity to "opt out" of the rider that funds PEC's DSM/EE programs. By foregoing the "opt out", customers become eligible for DSM/EE incentives, but commit to pay the rider for a period of 10 years.

Incentives. The program provides three types of participant incentives:

- **A one-time participation incentive of $50/demonstrated kW** -- intended to enhance customer acquisition and to support customer investment related to program participation, including purchase and installation of automated controls;
- **A monthly availability credit of $2.50/contracted kW** -- intended to provide steady payment streams and ensure readiness;
- **An event performance credit of $5/curtailed kW** -- intended to increase resource reliability through an emphasis on event compliance.

This three-part incentive structure was selected in order to benefit customers for responding to more events and to be less punitive than current curtailable rates where customers meet their commitments or pay stiff penalties. This structure also ensures that PEC pays for performance but limits its costs when few events are called. As a 'pay for play' program, it ensures that customers will receive more incentives when the need for peak reduction is high.

1.3 *Reported Program Participation and Savings*

Fourteen customers actively participated in PEC's DRA program in the summer of 2011, representing 31 unique sites and 39 meters. Nineteen of the 39 meters are at commercial sites and five are at governmental sites. Sixteen meters belong to a single manufacturing company— the very large industrial participant (VLIP)— at eight unique sites and the four remaining meters are at other industrial sites.

---

[4] In practice, events have in PY2010 and PY2011 only been called for six hours.

An overview of the participating customers is presented in Table 3, including number of meters and sites by customer type and the average demand reduction reported by PEC over the three 2011 events, by customer type.

Table 3: Summary of Participating Companies and Agencies

Sector	Customer Type	Number of Participants	Number of Sites	Number of Meters	Avg Reported Reduction per Meter (kW)*
Commercial	Warehouse/Distribution	1	1	1	654
Industrial	Manufacturing	4	14	20	308**
Governmental	Gov't Institution	1	2	2	205
Governmental	Water Treatment	3	3	3	522
Commercial	Grocery	2	8	8	271
Commercial	Data Center	1	1	3	261
Commercial	Office	1	1	1	200
Commercial	Hospital/Medical	1	1	1	2,620
	Total Program:	**14**	**31**	**39**	

**As reported by PEC, average over three events.*

*** Note that 16 of the 20 meters for this sector belong to the VLIP and collectively account for an average reported reduction of 3,383 kW per event.*

Source: PEC DRA program database

PY2011 average event curtailments at individual meters ranged from just below 100 kW to more than 1,300 kW, as shown in Figure 1. One participant had an average reported demand reduction of 2,620 kW. This is not displayed as a bar in the chart below to avoid distorting the scale, although it is reported in an embedded text box. In this chart meters are segregated by sector: commercial/ governmental and industrial. The map in Figure 2 illustrates the locations and relative magnitude of curtailments of all participants.

Figure 1: Reported Load Reductions (kW) by Meter

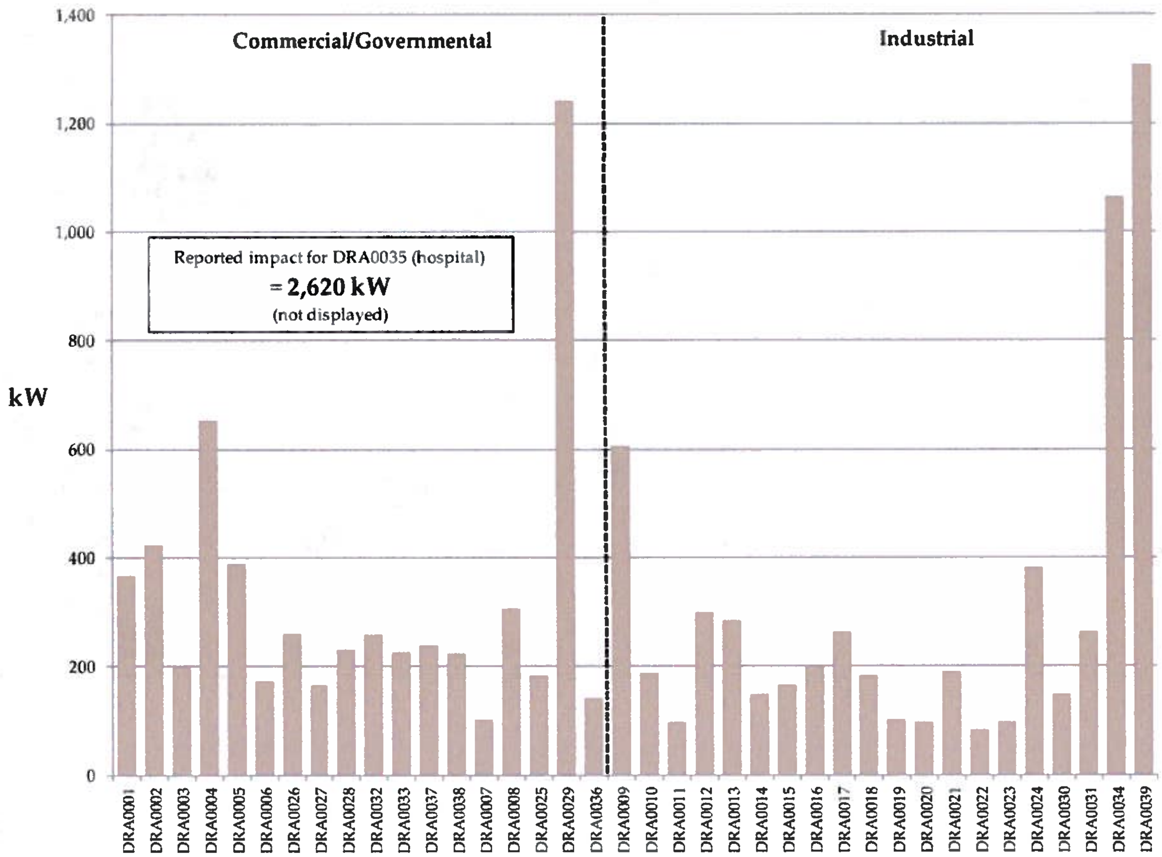
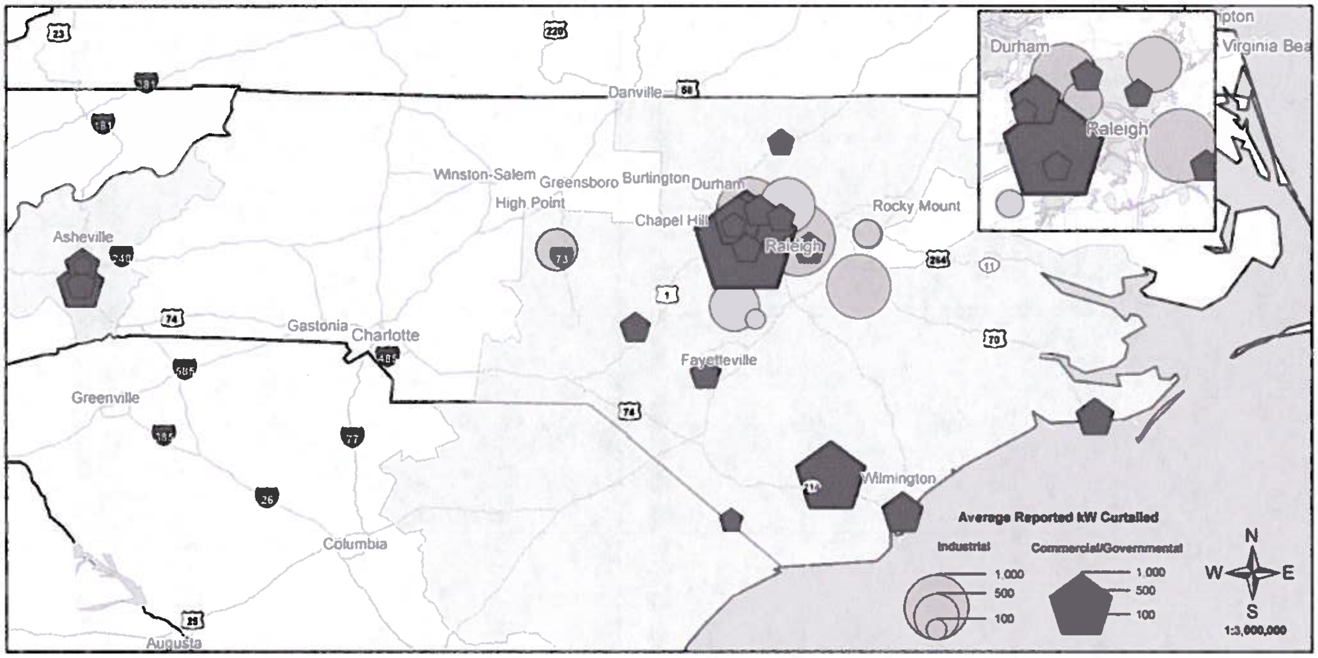


Figure 2: Locations of Participating Sites



Section 2. Evaluation Methods

This section describes the methods and data used by the evaluation team in estimating load impacts for PY 2011. No process evaluation was conducted for this program year. These methods address two processes: 1) replication of the savings calculations provided by PEC (referred to as the "reported" savings) and 2) a more data-intensive linear regression analysis to provide a more accurate estimate of the achieved load reductions (referred to as the "verified" impacts).

Estimating impacts of DR events is generally a matter of first estimating a "baseline" of what a customer's load would have been during the hours of the curtailment event if the event had not been called. Actual measured loads are then subtracted from this baseline to estimate load reductions. The baseline estimation methods used by PEC and by the evaluation team are discussed below.

The evaluation team used the following data in its analysis:

» Quarter hourly interval data for 39 DRA program participants between May 1st and September 30th, 2011.

» Twenty-minute interval temperature data (averaged to hourly frequency) from twelve National Oceanic and Atmospheric Administration (NOAA) weather stations.

» Event logs supplied by PEC indicating the dates and start and ending times of each event, as well as the time at which participants were notified of an imminent event.

Using this data, the evaluation team conducted three principal sets of analyses:

1. **Replication of the savings calculations provided by PEC**, which estimated baselines using the three "non-excluded" days immediately prior to an event (see below for details)
2. **Determination of the most accurate baseline estimation method for the VLIP** with 16 meters, and application of this method for estimating load reductions.
3. **Estimation of the impact of events for all meters** not belonging to the VLIP referenced above using a regression-based framework. In the PY 2010 evaluation, a regression with a symmetric four-hour day-of adjustment was determined to be the most accurate out of a menu of different day-matching and regression-based estimation methods for estimating participant baselines. Estimation of the impact of events for meters belonging to the VLIP mentioned above were calculated using the method determined to be most accurate in number 2, above.

Evaluations of DSM/EE programs commonly estimate a net-to-gross (NTG) ratio based on the evaluated percentage of demand reductions which may be ascribed either to free-ridership (which reduces the NTG ratio) or program spillover (which increases NTG). Free ridership is typically defined as the percentage of demand reductions that would have occurred anyway, absent the presence of the program. Spillover is typically defined as incremental demand reductions undertaken by a program's participants not directly incented or promoted by the program administrator.

In the case of demand response programs such as DRA, there is no reason to expect that a customer would curtail loads during the event periods (the timing of which would be unknown to the customer absent participation in the program). Furthermore, since demand reductions are estimated relative to an estimated baseline which captures expected participant behavior absent an event, the analysis inherently accounts for free ridership and spillover; that is, absent the DRA program, none of the observed demand reductions would have taken place. Based on the above considerations, the evaluation team considers the NTG ratio for this impact analysis of the DRA program to be 1.0[5].

2.1 *Replication of the PEC Savings Calculations*

PEC estimated load reductions using an internally developed baseline calculation method as described in *Demand Response Automation Rider DRA-1A (North Carolina), and DRA-2A (South Carolina)*[6] filed by PEC in June of 2010. The evaluation team replicated PEC's algorithm in order to confirm the results reported by PEC. The PEC algorithm involves the analysis of curtailment period hours to select the three immediate prior, non-excluded (holidays, weekends, and curtailment days) similar days. If any individual day is below 50% of the three-day average, the day is discarded as an outlier and the next prior non-excluded day is used. If there are not three qualifying days out of the ten non-excluded days prior to the event, the algorithm reverts to using the three most immediate non-excluded days prior to the event. The average demand over the three selected days during the hours corresponding to those in which the event was called on the event day is the baseline used to calculate impacts and participant incentive payments.[7]

---

[5] As noted below, the VLIP tends to have unpredictable demand, which can be either significantly higher or lower than average on any given day. On some days a customer with this pattern may be a free rider, but on others it may receive less credit than warranted by its verified reductions.

[6] North Carolina Rider, DRA-1A: https://www.progress-energy.com/assets/www/docs/company/NCRiderDRA.pdf
South Carolina Ricer, DRA-2A: https://www.progress-energy.com/assets/www/docs/business/scriderdra.pdf

[7] It should be noted that in PY 2011 one participant meter, DRA0039, was subject to an important exception to the evaluation procedure described above. The customer to whom this meter belonged is an industrial customer that was still partially under construction over the summer of 2011. This meant that there were working shifts in place only four days a week, with one day a week put aside for line-cleaning. In order to allow this customer to participate in the program, PEC, dropped known line-cleaning days from the sample when calculating this meter's baseline. PEC confirmed to Navigant that line-cleaning days that were dropped from the sample were July 7th, July 21st and

The details of the PEC algorithm are described in more detail in Appendix A of the PY 2010 report. For the PY 2010 report, the evaluation team, in addition to verifying the results obtained by PEC using the baseline algorithm, explored possible variations on that algorithm for all participating meters. For PY 2011, this analysis was repeated for only the VLIP possessing sixteen of the thirty-nine meters participating in PY 2011. Understanding the variations requires an understanding of the base algorithm, a more expansive explanation of which may be found in the appendices of the PY2010 report.

2.2 *Determination of the Most Accurate Baseline Estimation Method, Day-Matching Algorithms and Regression Baselines*

The fundamental method by which the most accurate baseline was determined for the VLIP, and the calculation of the day-matching algorithms and regression baselines is unchanged from PY2010. A thorough description of all of these techniques may be found in the body and appendices of that report.

August 8th of 2011. In replicating PEC's algorithm, Navigant also dropped these days from its sample when replicating the PEC algorithm.

Section 3. Program Impacts

This chapter describes the findings from the evaluation team's analysis of load reduction impacts for the DRA program for PY 2011. It is divided into four sections:

Section 3.1: **Confirmation of PEC's estimate of load reductions** (using the PEC settlement algorithm).

Section 3.2: **Accuracy assessment of alternative baseline estimation techniques**, applied specifically to the VLIP.

Section 3.3: **Verified demand impacts** using the baseline method found in the PY 2010 report to be the most accurate; for the VLIP, however, impacts were calculated using the baseline method found in 3.2 to be the most accurate for this customer.

Section 3.4 **Comparison of verified impacts with PEC's reported and contracted values**.

Three events were called during summer 2010, involving 39 unique customer meters. The EM&V analysis found average load reductions of 13.4 MW per event (roughly 345 kW per participant), or approximately 92% of the 14.6 MW figure reported by PEC in its DRA program database (Table 4).[8]

Table 4: Verified Load Reductions and EM&V Verification Rate

Load Reduction Category	July 12, 2011 (kW)	July 22, 2011 (kW)	Aug 08, 2011 (kW)	Avg. Total Reduction Over Three Events (kW)
Reported (PEC Database)	13,533	15,322	14,849	14,568
Verified	12,820	13,889	13,453	13,387
Realization Rate (Verified Reductions/Reported Reductions)	95%	91%	91%	92%

Source: PEC DRA program database and Navigant analysis

[8] As noted previously, "reported" impacts are those impacts calculated by PEC using the DRA baseline algorithm. Verified impacts are based on a regression baseline. Both sets of impacts are net values, implicitly assuming an NTG ratio of 1.0. See Section 2 – Evaluation Methods for further discussion.

Other significant findings of the impact evaluation, by topic areas, are as follows:

Approved Baseline Methodology

- **Finding #1:** PEC's method for calculating DRA baselines and estimating load impacts is both transparent and replicable and therefore appropriate for purposes of determining payments to participants.
- **Finding #2:** The EM&V confirmed the accuracy of PEC's estimates based on the approved methodology.

Alternative Baseline Methodologies Applied to the Very Large Industrial Participant

- **Finding #3:** There is little incremental difference in the accuracy of alternative baselines regardless of the number of days used in "day-matching" approaches.
- **Finding #4:** "Day-of adjustments" to baseline estimation do not produce a significant improvement in accuracy of baseline estimation for the VLIP. This is in contrast to dramatic improvement such adjustments offer for the estimation of the other participants' baselines, as noted in the PY2010 EM&V report of December 2011.
- **Finding #5:** Regression models using available data from all non-holiday, non-event weekdays provide a more accurate baseline, even without a day-of adjustment.
- **Finding #6:** The PEC algorithm, due to its exclusion of below-average "outlier" days, tends to greatly over-estimate the demand reduction impacts of DRA events on the VLIP with sixteen meters.

Verified Impacts

- **Finding #7:** Using the most accurate baseline method tested, the evaluation team verified that participants as a whole achieved an average of 13.4 MW of demand reduction, approximately 92% of that reported and 99% of that contracted.
- **Finding #8:** More than half of all verified impacts were provided by the largest five meters out of the total of 39 meters evaluated.
- **Finding #9**: The highly variable demand patterns of the VLIP mean that it is very hard to distinguish to what degree the periodic low observed loads are the result of DRA participation and to what degree they are the result of other operational imperatives. Estimates of this participant's meter-level impacts should be treated more cautiously than the impacts of other participant meters. The VLIP's impacts should be considered at the participant level where overall accuracy is likely to be much higher than at the meter or even site level.

3.1 *Confirmation of PEC's Estimate of Load Reductions*

As noted above, part of the task assigned to the evaluation team was to replicate the PEC algorithm to confirm the validity of the results reported by PEC. The EM&V team replicated PEC's method of baseline estimation (including the exception noted above in 2.1) and thus verified that PEC's calculations of baselines and impacts for events on which data exists in 2011 were correct to within less than 0.01%.[9]

3.2 *Accuracy of Alternative Baseline Estimation Algorithms and Models for Large Industrial Customer*

As discussed in the review of evaluation methods (Section 2), alternative baseline approaches were reviewed for this analysis (see Appendix B of the PY 2010 report for baseline parameters). Each baseline estimation model—including PEC's DRA approach, the alternative day-matching scenarios, and the regression model—was tested for accuracy by predicting loads on days that did not have curtailment events and measuring the accuracy of the predicted baseline load compared to the actual load. Goodness-of-fit between two daily load curves (predicted load versus actual load) can be determined by summing the squared differences for each hour. The analysis was conducted using the 59 non-event days between June 1 and August 31, 2011 on which the maximum temperature was above 85 degrees Fahrenheit.

Figure 4 ranks the average sum of squared differences (SSD) for each baseline method applied to predict the recorded loads of the VLIP on each of the 59 non-event days in 2011 with maximum daily temperature above 85 Fahrenheit degrees. Preceding it, Figure 3 is reproduced from the PY 2010 report. Figure 3 ranks the average sum of squared differences (SSD) for each baseline method applied when each method is applied to predicting the load of *all* participating meters on each of the 60 non-event days in 2010 with maximum daily temperature above 85 Fahrenheit degrees.

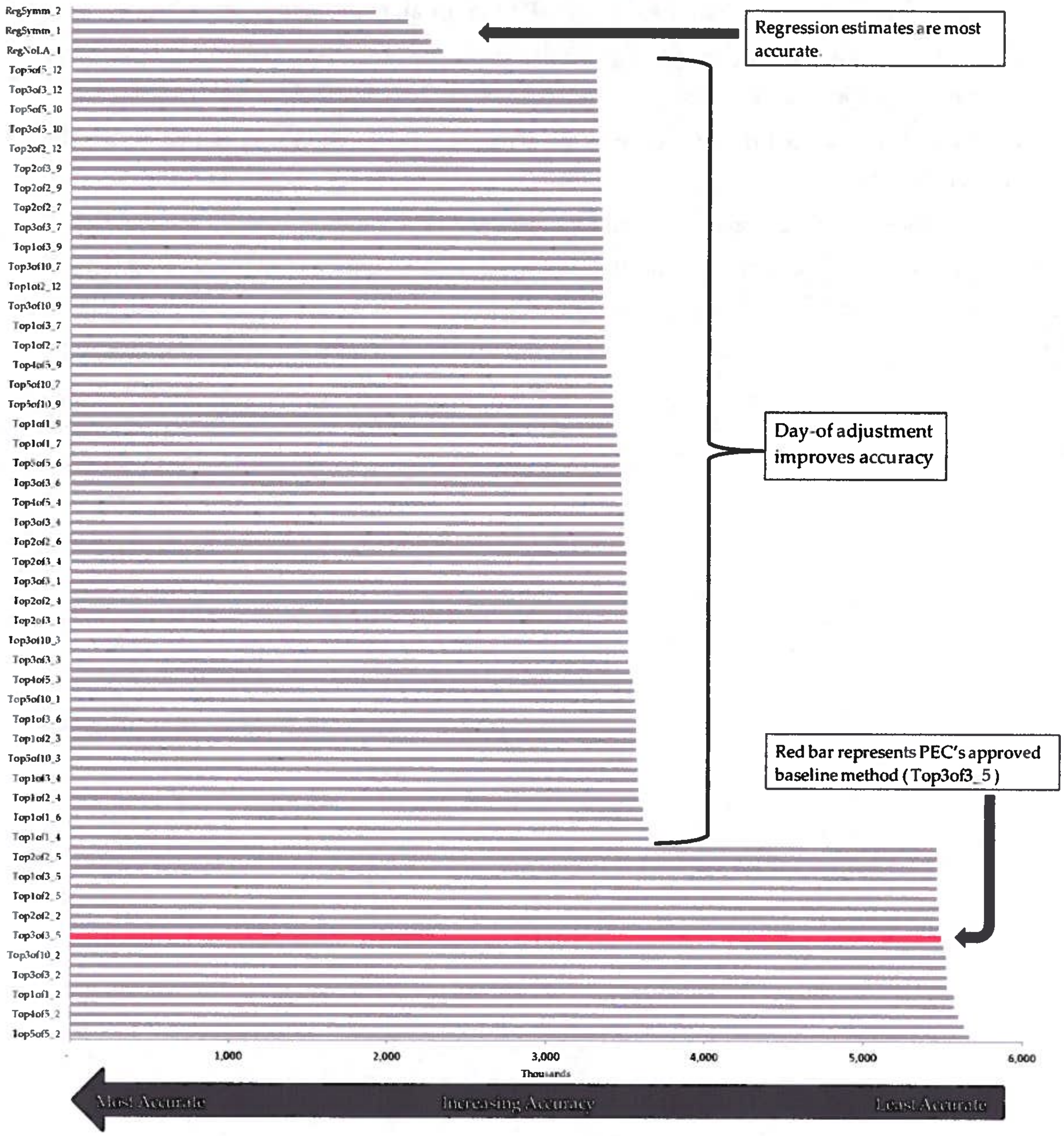
Figure 4 suggests several significant findings and is particularly interesting when contrasted with the same chart, Figure 3, ranking the methods across all participants, reproduced below from the PY 2010 report. Note that the chart from PY 2010's report resulted from the analysis on all participants at that time, whereas the chart below it is the result of analysis of only the sixteen meters belonging to the single large industrial customer in PY 2011. Some of the most significant findings:

---

[9] Note that this replication and verification does not include the replication of the demand reduction for meter DRA0023 on the July 12th event. In this case a meter component failure resulted in more than a week without interval meter data being recorded; the demand reduction reported for this meter on that day by PEC was simply the average demand reduction reported for the same meter in PY 2010. As insufficient data existed for this account to replicate the PEC algorithm for the 12th of July, Navigant has not replicated the PEC algorithm for this account on that date.

- PEC's 3-of-3 day-matching baseline approach (the red bar labeled with "Top3of3_5") is a less accurate predictor of the VLIP's loads than many alternative methods, however it performs significantly better for the VLIP than for all participants as a whole;
- A "day-of" adjustment does not lead to nearly so large an improvement in predictive power for the VLIP's meters as it does for the participants in general;
- There is relatively little difference in the accuracy of alternative day-matching baseline approaches;
- Regression methods provide additional accuracy, even without day-of adjustment.
- RegSymm_2 appears to deliver the most accurate estimates of DR impacts for both the VLIP and for DRA participants as a whole.

Figure 3: Accuracy Assessment of Alternative Baseline Methods – All Participants, Reproduced from PY2010 Report



Source: Navigant analysis

NAVIGANT

Figure 4: Accuracy Assessment of Alternative Baseline Methods – Only Very Large Industrial Customer, PY 2011

Source: Navigant analysis

The regression method with four hours' symmetric day-of load adjustment (represented by RegSymm_2) generates the least average SSD and therefore creates the most accurate baseline and most reliable aggregate load impact estimate for the large industrial participant's sixteen meters on high temperature weekdays. As noted above, this method also provides the most accurate baseline for participants as a whole. This finding is important in that DR events are often called on days with high temperatures that drive cooling demand and raise system loads. Consequently, **the EM&V team recommends use of the 4-hour symmetric baseline retroactively for estimation of system load impacts and prospectively for planning system resource needs.**

3.3 *Verified Impacts Estimated Using Regression-Based Baseline*

As described in section 3.2, the model found to most accurately predict demand for the large industrial participant with sixteen meters was the RegSymm_2 regression model. This is the same model that was found to most accurately predict demand for all the participants in PY 2010. This regression-based baseline also incorporated symmetric 4-hour day-of adjustment. All "verified" impacts discussed below are based on this model[10].

Three events were called during summer 2010, involving 39 unique customer meters. The EM&V analysis found average load reductions of 13.4 MW per event (roughly 345 kW per participant), or approximately 92% of the 14.6 MW figure reported by PEC in its DRA program database (Table 5).[11]

---

[10] Readers should note that two additional dummy variables were added to the model to take into account the idiosyncrasies of two of the meters evaluated. The first additional dummy variable was one which flagged the three known line-cleaning days for meter DRA0039 (see discussion in Section 3.1above). This was included to prevent observations on line-cleaning days from biasing the parameter estimates of the regression estimated for this meter. The second additional dummy variable was set to equal one when day *t* was a Friday and the meter was DRA0020. In the course of normal model diagnostics, the evaluation team found that this meter-site always had very low levels of consumption on Fridays compared to other days of the week and included the dummy variable to prevent observations for this customer on Fridays from biasing the parameter estimates of the regression estimated for this meter. Neither dummy variable in any way affected the results of any of the other meters.

[11] As noted previously, "reported" impacts are those impacts calculated by PEC using the DRA baseline algorithm. Verified impacts are net values, implicitly assuming an NTG ratio of 1.0. See Section 2 – Evaluation Methods

for further discussion.

Table 5: Verified Load Reductions and EM&V Verification Rate

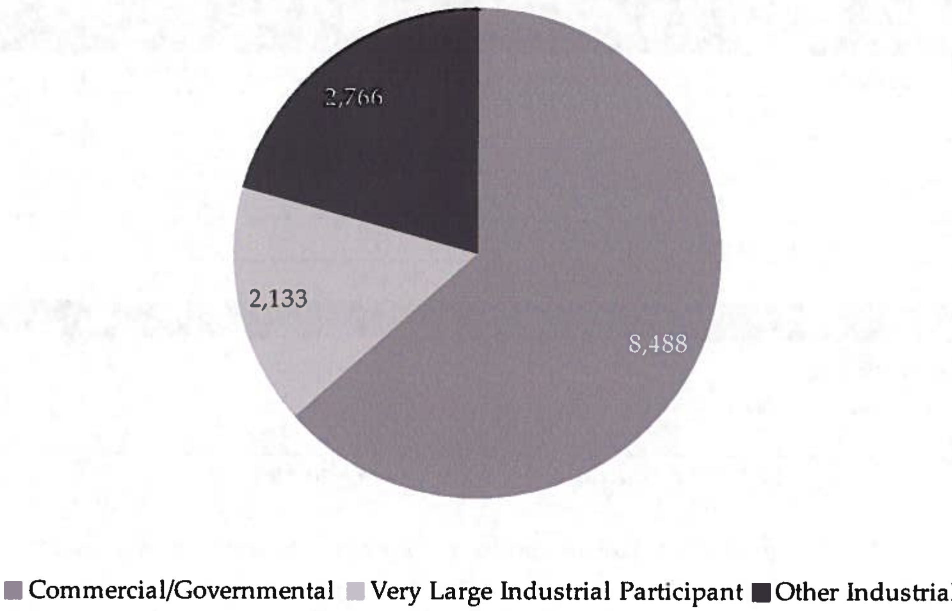
Load Reduction Category	July 12, 2011 (kW)	July 22, 2011 (kW)	Aug 08, 2011 (kW)	Avg. Total Reduction Over Three Events (kW)
Reported (PEC Database)	13,533	15,322	14,849	14,568
Verified				
Com/Gov't	*7,928*	*9,022*	*8,515*	*8,488*
VLIP	*2,190*	*2,149*	*2,059*	*2,133*
Other Ind.	*2,702*	*2,719*	*2,878*	*2,766*
Verified - Total	12,820	13,889	13,453	13,387
Verified Realization Rate (Verified Reductions/Reported Reductions)	95%	91%	91%	92%

Source: PEC DRA program database and Navigant analysis

For 2011 the EM&V team verified that the nineteen commercial/governmental meters realized a total of 8,488 kW of load reductions, which accounts for approximately 63% of the total kW reduction; the sixteen industrial meters belonging to the VLIP realized a total of 2,133 kW of load reductions[12], which accounts for approximately 16% of the total kW reduction. The balance of load reductions, 2,766 kW or 21% of the total, were made up by meters located at industrial sites not belonging to the very large industrial participant. This distribution is shown in Figure 5, below.

[12] For program year 2010, the same participant's meters realized load reductions of approximately 2,181 kW.

Figure 5: Share of Total Verified kW Reduction: Commercial vs. Industrial



Source: PEC DRA program database and Navigant analysis

The following discussion provides a summary of load impact findings based on a linear-regression baseline method identified by the EM&V team as the most accurate for predicting customers' loads. Load reductions were estimated for individual participants for each event. Verified program savings were then calculated as the average across each of the three 2011 events of the total reductions across all 39 participants' meters.

PEC had already estimated program impacts to be approximately 108% of the aggregate contracted load reductions, or 14.6 MW. The EM&V analysis verified roughly 92% of these reductions (13.4 MW). The average contracted, PEC-reported, and verified load curtailment for each participant meter is shown below in Table 6.

Table 6: Average Contracted, Reported and Verified Loads by Site

Comercial/Governmental				Industrial			
Participant Site	Contracted kW	PEC Reported kW	Verified kW	Participant Site	Contracted kW	PEC Reported kW	Verified kW
DRA0001	383	367	366	DRA0009	450	605	545
DRA0002	383	424	425	DRA0010	75	187	158
DRA0003	150	200	206	DRA0011	75	98	66
DRA0004	660	654	700	DRA0012	300	298	165
DRA0005	370	388	404	DRA0013	75	285	235
DRA0006	180	172	175	DRA0014	75	149	127
DRA0026	255	261	270	DRA0015	150	165	84
DRA0027 (1)	254	165	174	DRA0016	200	199	53
DRA0028	246	232	242	DRA0017	200	262	61
DRA0032	257	258	265	DRA0018	180	181	45
DRA0033	244	225	231	DRA0019	100	103	47
DRA0035	2,500	2,620	2,681	DRA0020	75	98	74
DRA0037	249	239	243	DRA0021	242	190	153
DRA0038	200	223	210	DRA0022	75	83	25
DRA0007	80	103	105	DRA0023 (2)	75	97	31
DRA0008	275	307	288	DRA0024	400	381	274
DRA0025 (1)	292	182	188	DRA0030	150	148	157
DRA0029	1,200	1,242	1,177	DRA0031	240	263	289
DRA0036	200	142	141	DRA0034	780	1,064	1,201
				DRA0039	1,250	1,308	1,119
Sub-Total	**8,377**	**8,402**	**8,488**	**Sub-Total**	**5,167**	**6,166**	**4,899 (3)**
				Sub-Total for Meters DRA0009 Through DRA0024 (VLIP) Only:			
					2,747	**3,383**	**2,133 (3)**

(1) Average includes one event for which contracted, reported and verified kW = 0

(2) Average calculated over only two of three events - no data for July 12th event

(3) Note that this sub-total reflects weighting applied to account for the meter for which data was missing for the first event and thus will not match the sum of the values in this column which are not weighted by the number of events for which data exists.

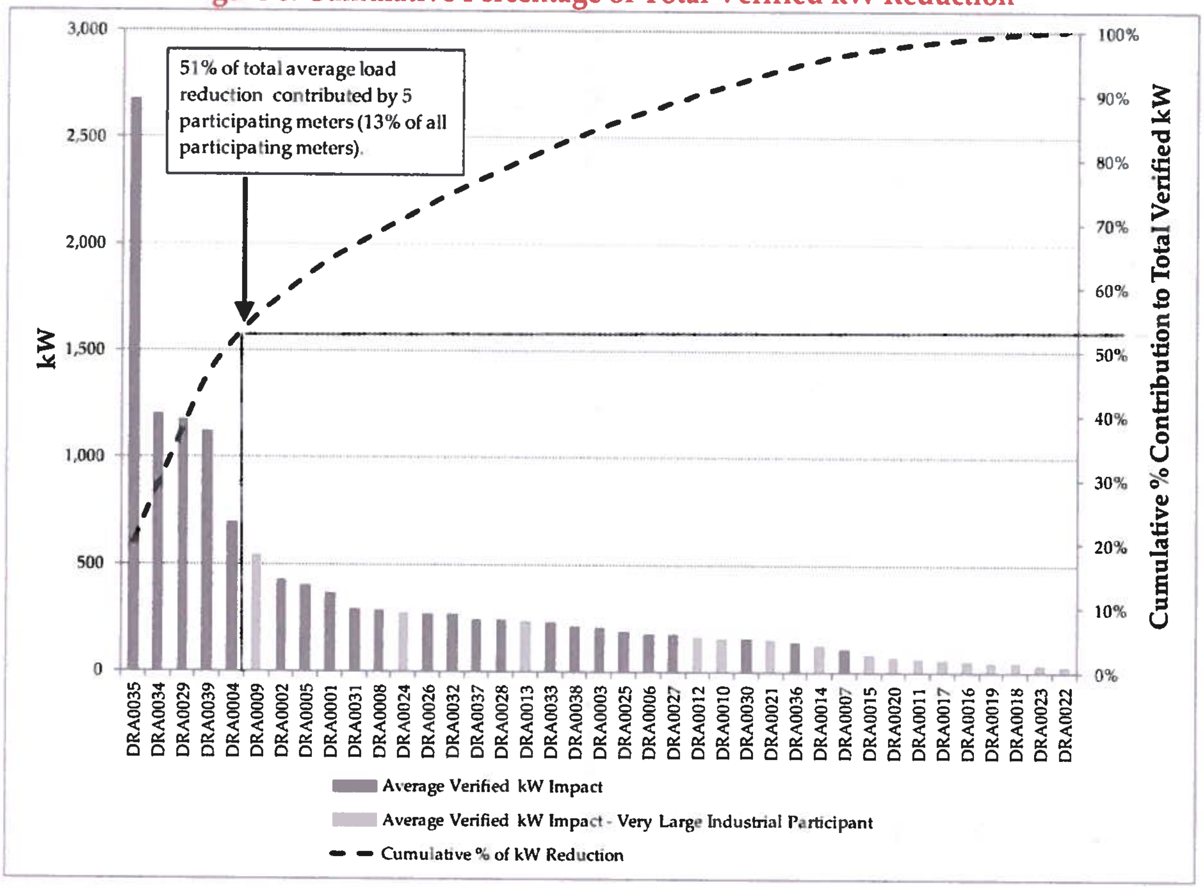
Source: PEC DRA program database and Navigant analysis

Verification rates at the portfolio level are driven by findings for individual meters. Five of the thirty-nine participating meters in 2011[13] account for more than half of all reductions and thus drive overall findings. Figure 6 ranks the meters by the amount of verified kW reduction in descending order, and it illustrates the decrease in load reductions between the largest and

[13] The five are composed of two commercial, one governmental and two industrial, none belonging to the VLIP.

smallest contributors in the program. It is revealing to note that the single largest contributing meter accounts for 20% of the overall verified reduction.

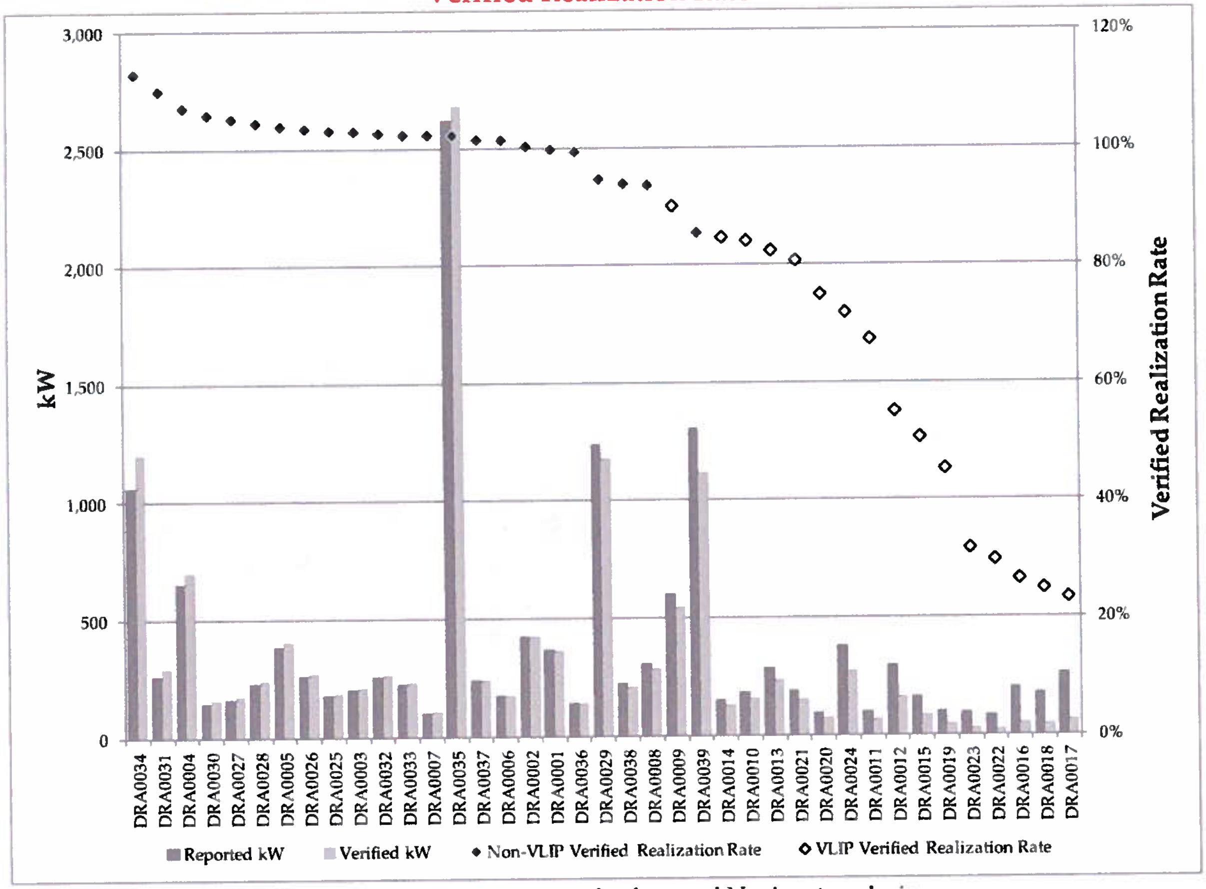
Figure 6: Cumulative Percentage of Total Verified kW Reduction



Source: PEC DRA program database and Navigant analysis

In Figure 6, it's clear that while a high proportion of the VLIP's average verified reductions are quite low, they are not overwhelmingly so – in fact one of the VLIP's meters has the fifth highest verified demand reduction. It is interesting therefore to re-examine these results by plotting the reported and verified demand reductions and verified realization rate (verified kW divided by reported kW) once they have been sorted by verified realization rate – see Figure 7, below. Abruptly, a pattern becomes clear – note that the black diamonds (right axis) in Figure 7 show the verified realization rate for the non-VLIP meters and the white diamonds show the verified realization rate for the VLIP meters.

Figure 7: Reported and Verified Demand Reductions, Verified Realization Rate, Sorted by Verified Realization Rate



Source: PEC DRA program database and Navigant analysis

While it appears that more than half of the meters have verified realization rates exceeding 95%, all of the meters associated with the VLIP have verified realization rates that are less than 100%, in some cases considerably less than one. Section 3.4.1, below, compares the impacts reported by PEC and those verified by the evaluation team and provides the results of the evaluation team's investigation of this skew in the VLIP's verified realization rates.

3.4 *Comparison of Verified Impacts and PEC Values – Reported and Contracted*

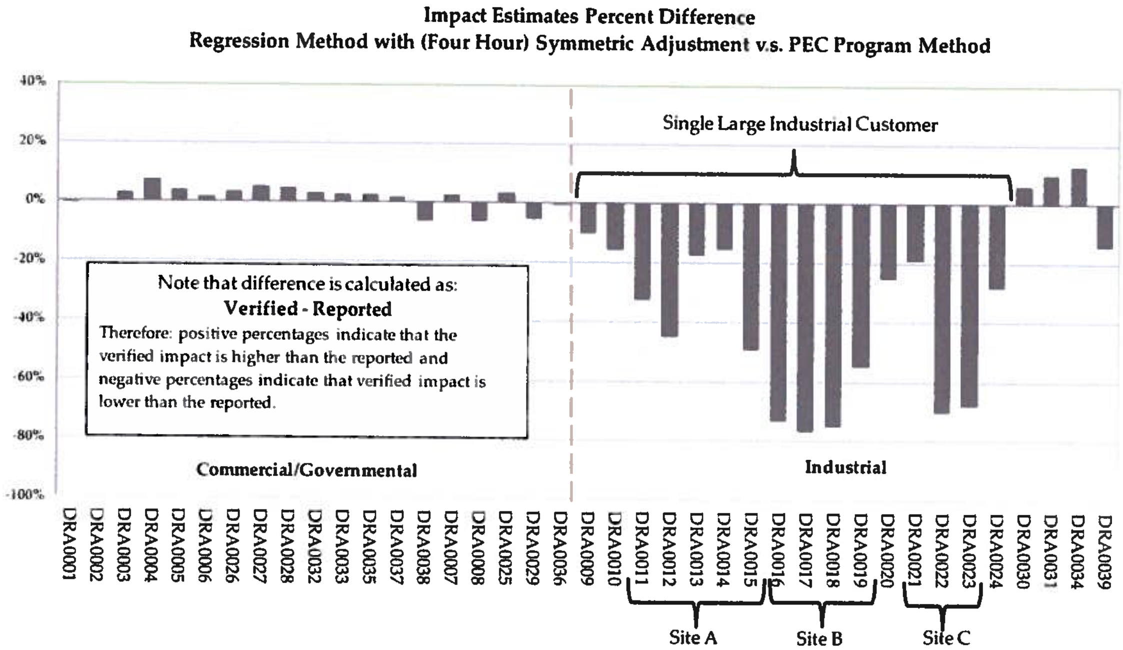
The two sections below compare:

- The verified demand impacts estimated by the evaluation team with those reported by PEC (calculated using its settlement algorithm) and
- The verified demand impacts estimated by the evaluation team with the contracted DR capacity of the participants.

3.4.1 Comparison of Verified and Reported Demand Reductions

As may be seen in Figure 8 the magnitude of the differences in load curtailment between impacts estimated using the PEC algorithm and the RegSymm_2 most reliable regression model were not consistent across all participant meters. In fact, it appears that in most cases, excluding the VLIP with sixteen meters, the PEC algorithm has underestimated the impact of DR. In the case of the VLIP, however, it is apparent that the PEC algorithm has overestimated the impact of DR.

Figure 8: Changes in Impact Estimates – Regression vs. PEC Program Method



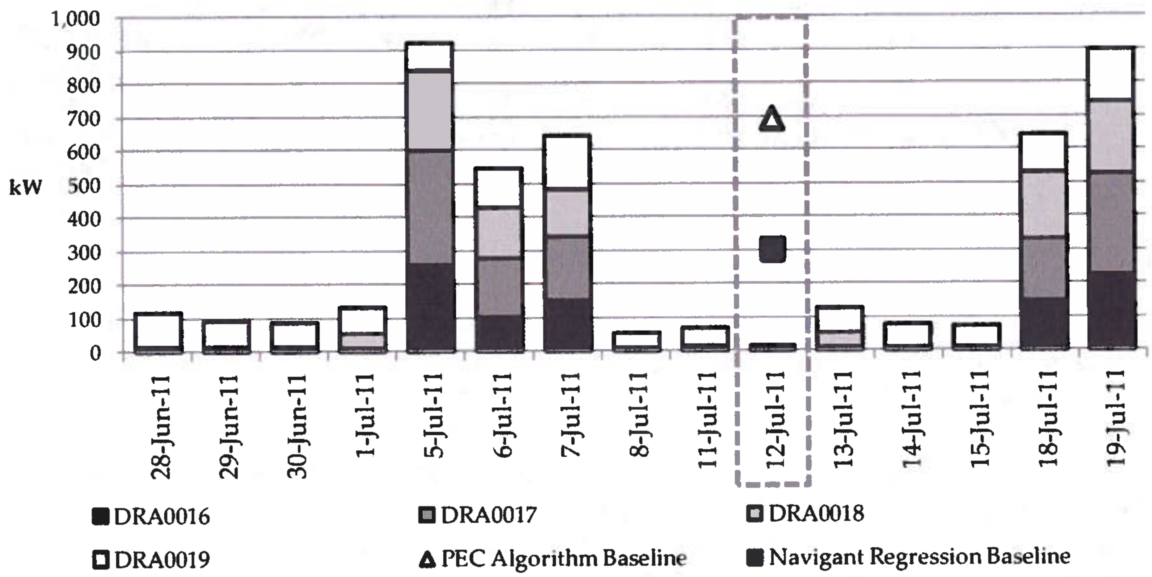
Source: PEC DRA program database and Navigant analysis

In the PY 2010 analysis, it was found that the PEC algorithm frequently overestimated the DR impacts observed on the sixteen meters of the only (at that time) industrial participant. The largest deviation of that year in fact, -78% for DRA0024, was larger than any of the deviations reported above in Figure 8. That said, the average absolute deviation observed across the large industrial customer's sixteen meters in PY 2010 was only 29%, compared with an absolute average deviation for the same customer of 42% in PY 2011.

These findings raise the question: Why does the PEC algorithm appear to be systematically over-estimating impacts? The evaluation team investigated the largest deviations more closely, concentrating its efforts on the meters located at the large industrial participant's Site B (DRA0016 through DRA0019, see Figure 8) since this set of meters had the highest deviation in absolute terms.

Investigation into this site's data revealed a highly variable pattern of demand, which served to confound the PEC algorithm and, to a lesser degree, the Navigant regression. See, for example Figure 9, below which shows the average demand on non-holiday weekdays during the hours when events were called over fifteen days for the various meters installed at Site B. The event day is enclosed in the dashed box, and Navigant and PEC's respective estimates of that day's baseline are indicated.

Figure 9: Average Demand (kW), Site B of Large Industrial Participant, 1p.m. Through 7p.m. EST, Non-Holiday Weekdays, June 28 Through July 19.

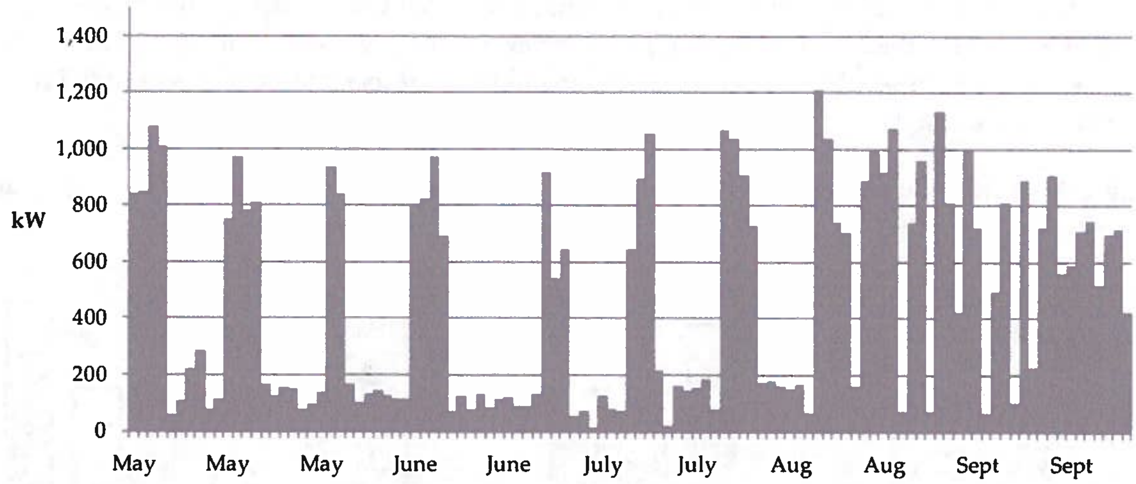


Source: PEC DRA program database and Navigant analysis

Although it is impossible to know with certainty, the pattern of demand certainly suggests that low levels of demand at this site on the event day were due not to the DRA program but some other operational imperative. Recall that all the dates shown in the chart above are on non-holiday weekdays and the demand reported is the average demand per meter during the same hours of the day as those in which events are called. It is also worth noting that the erratic pattern for this site is not isolated to the period shown in Figure 9, but appears to have been the norm over the entire summer period, gradually becoming more consistent in the later weeks of the sample period (i.e., September), see Figure 10, below. As previously, no holidays or weekends are included in this chart.

Note that the evaluation team investigated whether there was a pattern in the "not working" days and found that the average level of demand during working hours of non-holiday weekdays did not vary substantially by day of the week. That is, it does not appear as though "not working" days occur predominantly on (for example) Mondays or Tuesdays, but rather are spread relatively evenly across the five days of the working week.

Figure 10: Average Aggregate Demand (kW), Site B of Large Industrial Participant, 1p.m. through 7p.m. EST, Non-Holiday Weekdays, Entire Sample Period.



Source: PEC DRA program database and Navigant analysis

If we accept the working hypothesis that low levels of demand on the event day are due to some factor other than the DRA program, then why has the PEC algorithm so grossly over-estimated the baseline?

This is a function of the PEC algorithm's qualifying rules with regard to outliers. Recall that when a group of three days is selected to calculate the baseline, if any of those three days has an average level of demand that is less than half of the average level of demand of all three days together, it is dropped from the sample. Thus, in the example above, both July 8th and July 11th (neither of which is more than half of the average event period consumption across July 7th, 8th and 11th) will be dropped from the calculation, leading to what appears to be an over-estimated baseline. It is noteworthy that of the various types of baseline estimation techniques tested above in Figure 4, nearly two-thirds (65%) of those in the most accurate half *did not* drop outliers (as the PEC algorithm does) whereas slightly more than two-thirds (71%) of those in the least accurate half dropped outliers. This suggests that dropping outliers tends to reduce the accuracy of the baseline estimate for this particular participant.

Under the working hypothesis established above (that demand reduction is due to operational imperatives and not the effect of the DRA program), the regression baseline is also clearly too high, although much less so than the PEC algorithm. How problematic is this when using these results for planning purposes? The evaluation team believes that the expected value of the regression baseline method across multiple events and meters is a reasonable estimate, even if the method does not produce accurate results for each individual meter for each event.

Recall that the regression baseline is calculated first by estimating the relationship between a given meter-site's demand and weather, and between a given meter-site's demand and the hour of the day. The baseline is calculated by applying these estimated relationships to the actual weather and time-of-day values observed during the event. In fact there are really two sets of relationships, or regimes, at work: the relationship that exists on a day when the site is "operational" and the relationship that exists on a day when the site is "not operational" (i.e., the manufacturing process is not in use). The single relationship calculated by the regression is therefore an average between the two regimes outlined above, weighted by the frequency with which the two regimes ("operational" and "not operational") occur. What this effectively means is that the regression-estimated baseline will be too low when the event day is "operational" and too high when the event day is "not operational" On average therefore, the expected value of the estimated demand impact will still be reasonably accurate. Put another way, while individual event- and meter-specific impact estimates may not be accurate, at the aggregate level, the overall estimate of the VLIP's demand reductions should be reasonably accurate. This is borne out by the proximity of this year's aggregate demand reduction impact for the VLIP (2,133 kW) and last year's (2,181 kW).

Unfortunately, it is impossible to explicitly control for the "not operational" days without some *a priori* knowledge about when and why they occur. While it seems reasonably clear that in Figure 9, July 8th is, for example, not an "operational day", the evaluation team has no basis to claim that the event day July 12th wasn't an "operational day" as opposed to simply the participant's reaction to the DRA event.[14]

The EM&V team is *not* proposing the regression method for settlement purposes (for verifying customer compliance for purposes of incentive payments). As noted earlier, settlement baselines should be easily replicable and able to be calculated within hours or days of an event. Conversely, regression analysis requires econometric modeling or customized software and meter data from an extended period of time, such as the summer months, and therefore is not appropriate for settlement purposes.

---

[14] The variable demand patterns of the VLIP, in combination with its average daily load profile (when a site is "operational") and the early hours of notification also explain why the day-of load adjustment had a much smaller impact on the accuracy of the various algorithms for the VLIP in PY2011 than it did for participants as a whole. The VLIP sites' loads are very much coincident with typical business hours – in the early hours of the morning the VLIP's sites are not in operation (on both "operational" and "non-operational" days). That is, whether a day is "operational" or not for a VLIP meter, that meter's demand at (for example) 6am will be very low. When notification is provided at 9am, the day-of adjustment compares the average demand on the event day in the two or four hours previous to notification with the predicted baseline demand in the same hours. Since, regardless of whether the meter site is "operational", this demand will be very low in those very early hours, the day-of adjustment does not correct for the fact that, on a given day, the meter-site may not be "operational". If notification were provided later in the day then the two- or four-hour window used by the day-of adjustment would be much more likely to detect if the day in question is "operational" or not.

PEC may wish to reconsider, however, the manner in which its algorithm treats outliers, perhaps by excluding days from the calculation based on some threshold of symmetric deviation in, for example the first ten non-holiday weekdays preceding the event. That is, dropping not just very low values (relative to the mean over some period long enough to avoid distortion by extreme individual days) but also the very high ones.

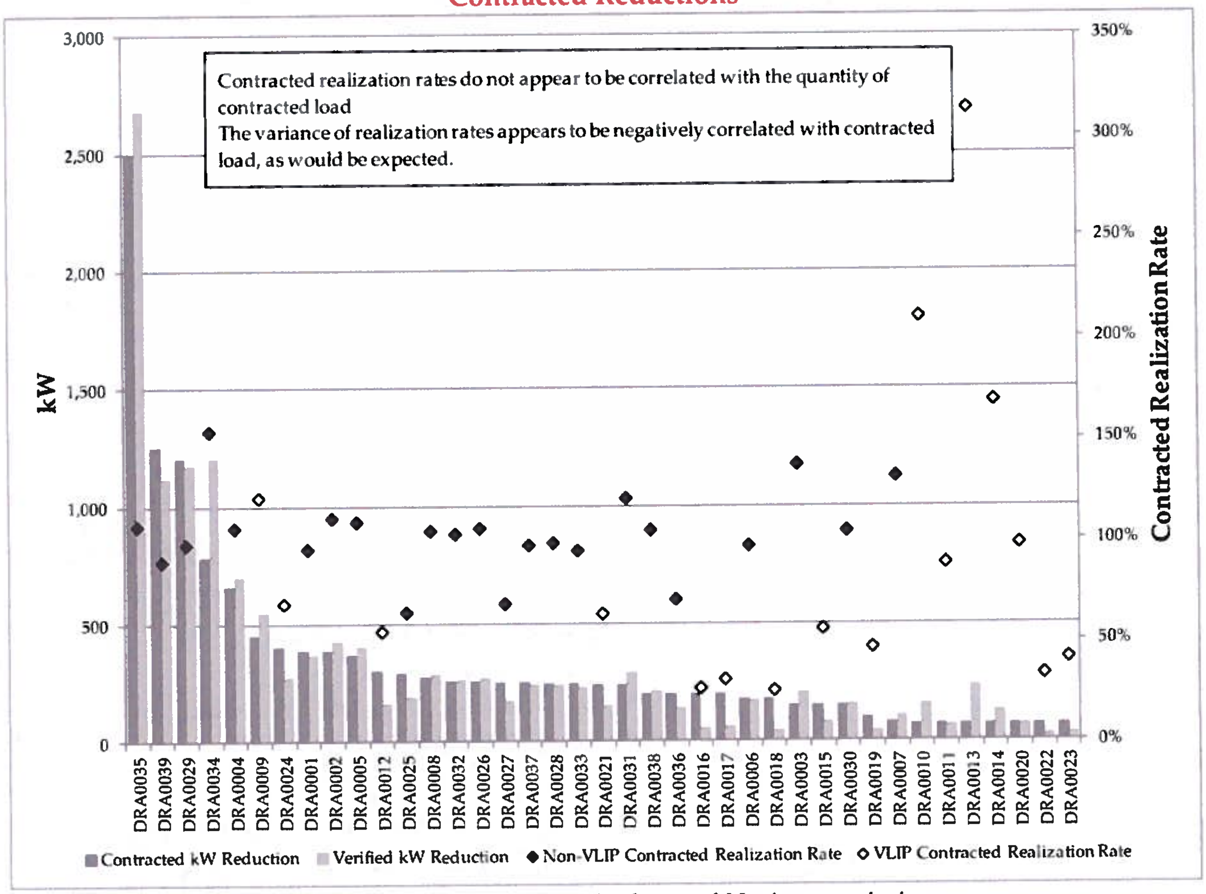
Another conclusion may also be drawn from the discussion above: that the VLIP's demand has become considerably more variable than it was in PY2010. In the evaluation of PY2010, it was found that the PEC algorithm sometimes underestimated demand reductions and sometimes over-estimated them (see Figure 3 of the PY2010 report). In contrast, for PY2011, the PEC algorithm has consistently overestimated the demand reduction for all of the VLIP's meters. Since the PEC algorithm will nearly always over-estimate demand reductions when there are a significant number of "non-operational" periods (due to the outlier exclusion rules) this suggests that for some of the VLIP meters at least, demand in PY2010 was much more consistent than in PY2011.

3.4.2 Comparison of Verified and Contracted Reductions

The evaluation's team principal mandate in the EM&V process is to provide verified demand impacts for all participants and report the verified realization rates (that is, verified kW divided by reported kW). The evaluation team, however, believes that it may be helpful for system planning purposes to provide some analysis of the contracted realization rates (contracted kW divided by verified kW). In this way, projections of program growth (through future contracted DR capability) may be applied to system planning directly, without first estimating what the corresponding reported DR impacts will be. **The discussion below discusses the contracted realization rates achieved in PY2011 with a particular emphasis on the behavior of the contracted realization rates of the VLIP.**

Figure 11, below, displays both contracted and verified reductions for each site, sorted from highest to lowest contracted amount. The top five meters from this figure correspond to the top five meters in Figure 6, above (although in a different order). From the black and white diamonds—which represent the contracted realization rate (verified reductions divided by contracted reductions, as seen on the right axis) for non-VLIP and VLIP meters respectively —it appears that the contracted realization rates for these top five meters hovers around 100%, with a low of roughly 90% and a high of nearly 154%. There is no apparent correlation between the quantity of load contracted for DR and the verified realization rate achieved by each meter. For example, note the meter site with one of the lowest amounts of load contracted (DRA0013) actually achieved a contracted realization rate of over 300% whereas a meter site with more than four times as much contracted load (DRA0024) achieved only a 69% contracted realization rate. There is, however, what appears to be a correlation between the variance of the contracted realization rates and the contracted load. This is to be expected. Since the contracted realization rate is the verified reduction divided by the contracted reduction, a given level difference between the two will result in a larger change in contracted realization rate for the one with the smaller denominator (contracted load) than the one with the larger denominator.

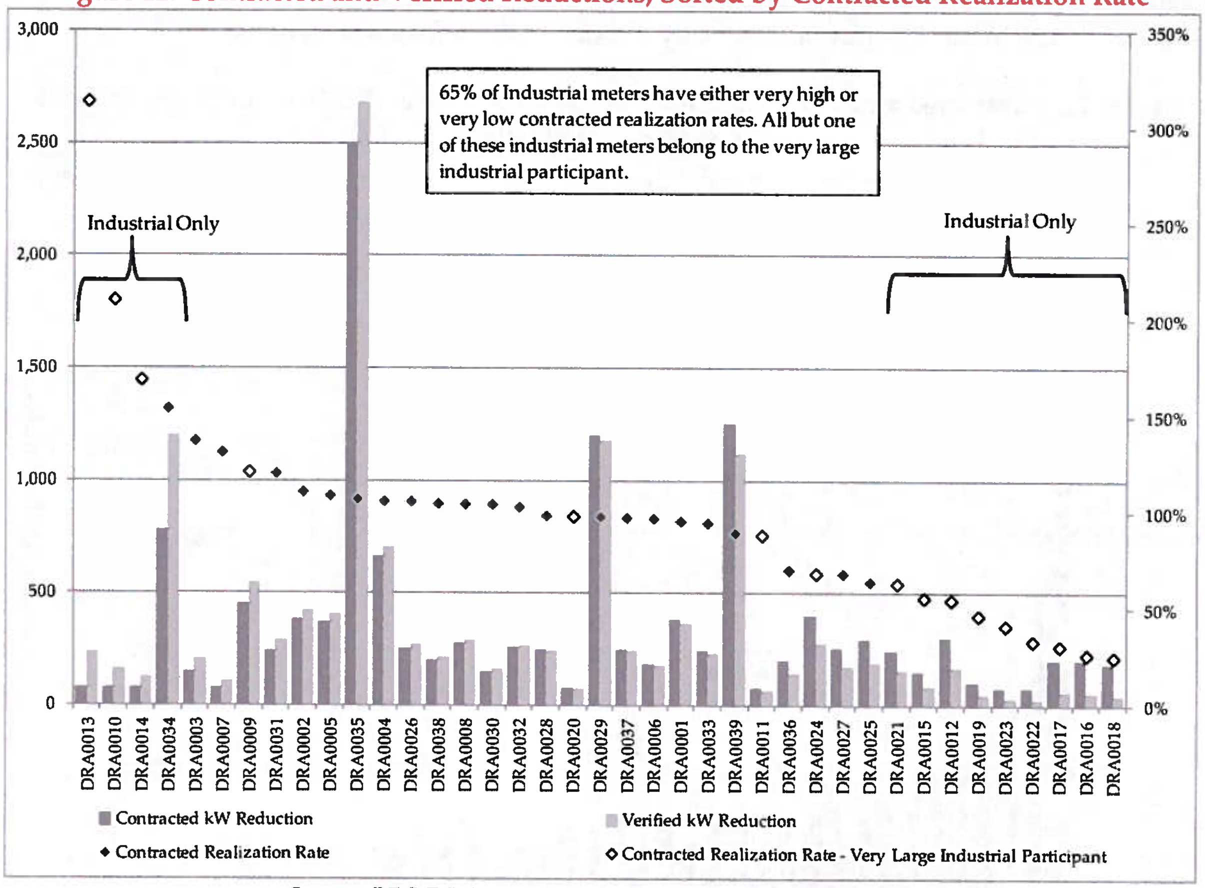
Figure 11: Contracted and Verified Reductions and Contracted Realization Rates, Sorted by Contracted Reductions



Source: PEC DRA program database and Navigant analysis

Although the contracted realization rate is uncorrelated with contracted load reduction another relationship becomes apparent when the meters are sorted by the contracted realization rate, as in Figure 12, below. Recall that in PY 2010, the evaluation team found that all but one of the meters with the highest contracted realization rates were at commercial/governmental sites and that all but one of the meters with the lowest contracted realization rates were at industrial sites. Interestingly, in PY 2011, the evaluation team now finds that while the lowest nine contracted realization rates all belong to industrial meters, so too do the highest four contracted realization rates. All but one of the thirteen most extreme contracted realization rates mentioned above belong to the meters of the VLIP (this participant has 16 meters in total). This contrast with the PY2010 values is another indication that the VLIP's demand has become much more variable in PY2011 than it was in PY2011.

Figure 12: Contracted and Verified Reductions, Sorted by Contracted Realization Rate

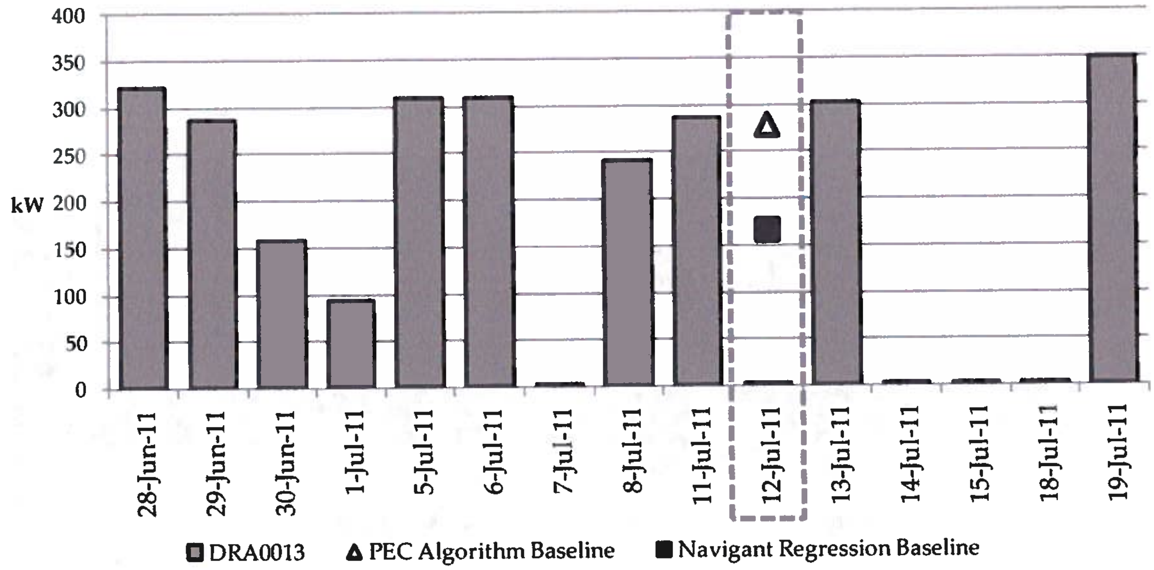


Source: PEC DRA program database and Navigant analysis

A possible explanation for this unexpected distribution of extreme contracted realization rates is that the variable demand behavior of the large industrial customer (discussed above in the context of one of this customer's sites – Site B – with four meters) is resulting in DRA demand reductions being attributed to site shut-downs that are unrelated to the DRA program in some cases and under-estimates of actual demand reductions attributable to the DRA program in others.

Investigation into the individual data series of the participant meter with the highest contracted realization rate (DRA0013) suggests that this may be a reasonable hypothesis but not nearly so conclusively as the example for Site B shown in above in Figure 9. Figure 13 shows the average demand during the same hours as the DRA events, on 15 non-holiday weekdays for the meter with the highest contracted realization rate (DRA0013, belonging to the VLIP) from June 28 through July 19. The PEC algorithm-estimated and regression-estimated baselines are also shown.

Figure 13: Average Demand (kW), Meter With Highest Contracted Realization Rate (DRA0013), 1p.m. Through 7p.m. EST, Non-Holiday Weekdays, June 28 Through July 19.

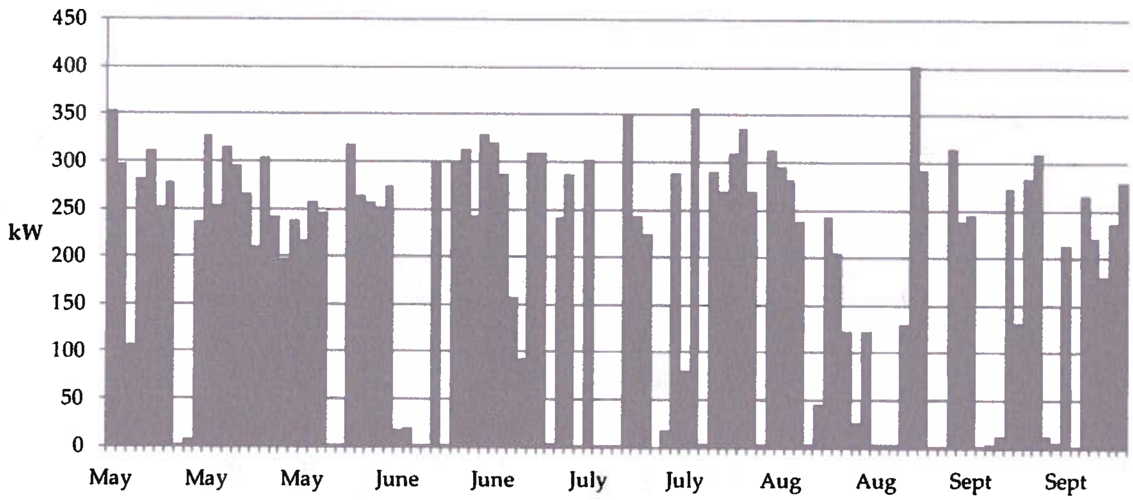


Source: PEC DRA program database and Navigant analysis

For Site B, (see Figure 9) it seems highly likely, given the large amount of days with virtually no consumption surrounding it, that the level of demand on the July 12 event day was due to some factor other than the DRA program. In the case above, for meter site DRA0013 (Figure 13), it is possible that the meter-site was not "operational" (due to factors other than the DRA program) on July 12 but it is equally likely that the meter-site is curtailing demand due to the DRA program incentives.

Note that both the day before and the day after the event day (indicated by the dashed box in Figure 13), the meter-site DRA0013 has a level of demand consistent with that occurring in the nine days prior to the event. Additionally, as seen in Figure 10 (above) showing the average demand for Site B, and Figure 14 (below) showing the average demand for meter DRA0013 when the VLIP shuts down a site, it does so for more than a single day at a time. This would suggest that the behavior seen in Figure 13 above, is motivated by DRA participation rather than some other operational imperative. It is also noteworthy that if in fact in this case the participant is responding to a DRA event rather than shutting down for some other reason, the PEC algorithm appears qualitatively to be more accurate than the regression. This reinforces the point that the VLIP's results should really be examined at the aggregate participant, and not meter, level.

Figure 14: Average Demand (kW), Meter With Highest Contracted Realization Rate (DRA0013),, 1p.m. through 7p.m. EST, Non-Holiday Weekdays, Entire Sample Period.



Source: PEC DRA program database and Navigant analysis

NAVIGANT

Section 4. Conclusions and Recommendations

The findings below highlight successes as well as specific recommendations to improve program design, operations, and outcomes.

4.1 *Conclusions*

The DRA program successfully ramped up operations and gained participation in 2011. Specific participation and impact findings are presented above in Section 3. Broader conclusions about the program's accomplishments and operations are as follows:

- **The DRA program has shown strong growth since PY2010**, nearly tripling the verified demand reductions and increasing the number of participating meters by 50%. The most significant participant addition in PY2011 was that of a hospital site which made the single largest contribution of any site and offered roughly the same demand reduction contribution as the program's largest customer.

- **The majority of demand reductions were made by a relatively small number of participants.** In fact, just three participating meters in PY2011 contributed more verified demand reduction to the program than was achieved in total in PY2010. This significant achievement emphasizes the importance of PEC account managers focusing recruitment on "high value" DRA participants capable of delivering large amounts of DR reliably.

- **In PY2011, the program achieved a more balanced distribution of participants across sectors relative to the prior year.** In PY2010, the program had only a single industrial participant (the VLIP); in PY2011, PEC added three other industrial participants. This diversity of participating sectors will tend to mean a more stable demand reduction capability for the program – swings in the economy that affect one sector in particular will have less of an impact on the program's capability to deliver reliable DR.

- A more in-depth analysis of the VLIP indicates that its demand has become more variable than in PY2010. **This participant's less predictable demand patterns mean that the PEC settlement algorithm, as a result of its treatment of outliers, is often overestimating the demand reductions.** Some recommendations for remedying this may be found in the section that follows. Additionally, the evaluation found that the use of a day-of adjustment to baseline calculations does not appear to significantly improve baseline estimates for this participant. This is because of the average daily load profile of this participant (very little demand before the start of standard business hours), and the fact that event notification occurs relatively early in the day.

4.2 *Recommendations*

The evaluation team recommends a variety of discrete actions for improving the results of the DRA program in future years. These are based on insights gleaned from participant interval data analyzed for PY2011, and a comparison with the results in PY2010.

Improvements to Baseline Estimation

1. **Modify the settlement baseline algorithm to exclude both low and high outlier days.** Currently, if a given day being considered for inclusion in the baseline algorithm has demand that is less than 50% of the average demand of all of the days being considered for inclusion, it is excluded from the calculation. The evaluation team recommends that PEC either: a) apply a parallel structure (also excluding days that have demand that are twice the average of all days being considered for inclusion), b) drop the outlier exclusion rule, or c) adopt an alternative rule to account for step changes in demand from day to day (such as those observed for the VLIP).

2. **Modify the settlement baseline algorithm to include a symmetric day-of adjustment.** The day-of adjustment was shown in PY2010 to considerably improve the accuracy of baseline estimation. Although analysis of the VLIP has shown that in that participant's case the day-of adjustment offers little improvement in accuracy, the adjustment would yield a more accurate baseline across all participants. As noted in the PY2010 report, adding a day-of adjustment would limit customers' ability to view their baseline loads on the day of an event and to make any necessary adjustments to their load curtailments. The EM&V assessment for PY2012 could assess the importance of this issue by surveying customers regarding how they respond to event notifications, and specifically whether they view or calculate their baseline loads prior to taking curtailment measures.

3. **Continue to use a regression-derived baseline incorporating a 4-hour symmetric day-of adjustment for the estimation of system load impacts and for planning system resource needs.** This method has been shown (in the PY2010 report) to be the most accurate baseline estimation method for participants as a whole, and in PY2011 has been shown to also be the most accurate baseline estimation method for the VLIP.

Change to Event Notification

4. **Consider leaving a smaller gap between event notification and event starting hours to improve the performance of the day-of load adjustment.** As noted above in reference to the VLIP, if a participant's demand is always the same in the hours of the day which typically precede notification, but frequently quite different in the hours of the day that follow the typical hour of notification, the day-of load adjustment will not greatly improve baseline accuracy. Naturally, in contemplating reducing the gap between event notification and the event itself, PEC must consider the degree to which such a change

could affect program participation and, ultimately, the DRA program's load curtailment capability.

Participant Recruitment

5. **Consider the consistency of potential participants' daily demand when targeting recruitment efforts and communicating with participants.** PEC account executives should be familiar with daily load patterns of prospective participants prior to engaging with the customers, as a customer with highly variable loads is likely to have less predictability with respect to its calculated baseline. This could result in overpayment for some events and underpayment for others. At a minimum, such customers should be aware that their unique load profile could result in payments that do not always match their actual event-specific reductions.

6. **Continue and enhance recruitment of new participants from the hospital sector.** The hospital currently enrolled in the DRA program contributed approximately 20% of the entire verified load reductions. If this capability is common amongst hospitals, further recruitment within this sector could greatly expand the DRA program's capabilities with relatively little impact on administrative requirements.